Children's social referencing reflects sensitivity to graded uncertainty

Emily Hembacher
ehembach@stanford.edu
Department of Psychology
Stanford University

Benjamin deMayo
bedemayo@stanford.edu
Department of Psychology
Stanford University

Michael C. Frank
mcfrank@stanford.edu
Department of Psychology
Stanford University

Abstract

The ability to monitor epistemic uncertainty is critical for self-directed learning. However, we still know little about young children's ability to detect uncertainty in their mental representations. Here we asked whether a spontaneous information gathering behavior – social referencing – is driven by uncertainty during early childhood. Children ages 2-5 completed a word-learning task in which they were presented with one or two objects, heard a label, and were asked to put the labeled object in a bucket. Referential ambiguity was manipulated through the number of objects present and their familiarity. In Experiment 1, when there were two novel objects and a novel label, the referent was ambiguous; when there were two familiar objects, or only one novel or familiar object, the referent was known or could be inferred. In Experiment 2, there were either two novel objects, two familiar objects, or one familiar and one novel object; in the latter case the referent could be inferred by excluding the familiar object. To further manipulate the availability of referential cues, the experimenter gazed at either the target or the center of the table while labeling the object. In both experiments, children looked at the experimenter more often while making their response when the referent was ambiguous. In Experiment 2, children also looked at the experimenter more when there was one familiar and one novel object, but only when the experimenter's gaze during labeling was uninformative. These results suggest that children's social referencing is a sensitive index of graded epistemic uncertainty.

Keywords: social referencing; help seeking; word learning; uncertainty.

Preschoolers quickly learn new concepts, rules, and language. They also actively explore and ask questions in ways that seem targeted to maximize learning (Chouinard, Harris, & Maratsos, 2007; Schulz & Bonawitz, 2007). However, we still have an incomplete understanding of young children's ability to monitor their own mental states, in particular, their epistemic uncertainty (Schneider, 2008; Sodian, Thoermer, Kristen, & Perst, 2012). Do preschool-aged children monitor uncertainty and guide their learning behaviors on the basis of this monitoring, or is early learning better characterized as a process of integrating information that is largely generated externally, for example, by social partners who act as teachers (Csibra & Gergely, 2006)?

A hallmark of successful uncertainty monitoring is being less confident when the probability of accuracy is lower (Robinson, Johnson, & Herndon, 1997). This includes awareness of complete ignorance, but also of graded evidence in mental representations (Ghetti, Lyons, Lazzarin, & Cornoldi, 2008). There is mixed evidence about whether young children can introspect on their mental states. For example, preschool-aged children fail at reporting on their ongoing thoughts (J. H. Flavell, Green, Flavell, Harris, & Astington, 1995), and 3-year-olds report being equally confident about correct and incorrect responses in a memory task (Hembacher & Ghetti, 2014). However, other studies have shown that preschoolers can distinguish between their own correct and incorrect answers in other tasks, although they are typically overconfident (Coughlin, Hembacher, Lyons, & Ghetti, 2015; Lyons & Ghetti, 2013; Paulus, Proust, & Sodian, 2013).

There is also protracted development throughout childhood of other metacognitive abilities. For example, children gradually improve in their ability to estimate future performance (Destan, Hembacher, Ghetti, & Roebers, 2014; Lipowski, Merriman, & Dunlosky, 2013) and selectively allocate learning time to difficult or less-well-learned materials (Metcalfe & Finn, 2013). However, most of these studies have relied upon explicit reports of uncertainty or learning progress, for example, by asking children to use a confidence scale. Perhaps children learn to respond appropriately to uncertainty in everyday learning situations before they can bring it fully into consciousness and report on it.

Several studies have provided evidence that children's spontaneous information-seeking behaviors might track uncertainty. Call and Carpenter (2001) had 2-year-olds choose between several tubes to find a hidden sticker. They found that the toddlers were more likely to peek inside a tube before choosing when they had not seen the baiting of the tubes compared to when they had, suggesting they were aware of their ignorance and managed to delay their response until they had a good answer. In another study, Goupil, Romand-Monnier, & Kouider (2016) found that 20-month-olds were more likely to seek help by looking at their parents when they were unable to respond accurately in a memory task. Thus, spontaneous information-gathering behaviors may provide a window into early uncertainty monitoring, and allow us to ask questions about it's development.

Here, we focus on the role of uncertainty in guiding social referencing – one form of information gathering – during word learning. Referencing a social partner can provide several forms of disambiguating information. First, children can follow a speaker's gaze direction to infer the referent of a new word, as people tend to look at objects they are referring to. By the second year of life infants follow a speaker's gaze and map labels to objects on the basis of gaze direction (D. A. Baldwin, 1991). There is also evidence that infants' propensity for gaze-following predicts later language development (Carpenter, Nagel, Tomasello, Butterworth, & Moore, 1998), highlighting the importance of this behavior for learning. In addition to monitoring gaze direction, people may reference a social partner's emotional reaction to a stimulus or event, which can help disambiguate the appropriateness of a response [ref]. Finally, looking at a social partner

can be taken as a bid for help [ref].

Social referencing can be an efficient source of disambiguating information, but is it driven by uncertainty during early childhood? It could be that social referencing is relatively cheap in terms of effort and time, so selectivity may be unnecessary. Similarly, children may simply respond to external events, for example, by fixating towards someone who is speaking, without monitoring their own knowledge states. Vaish, Demir and Baldwin (2011) addressed this question with infants. Thirteen to 18-month-olds sat across from an experimenter who produced a label (e.g., "a modi!) in the presence of one or two objects. They found that infants looked up to the experimenter more often when there were two objects present, and the referent was thus ambiguous. Thus, there is initial evidence that infants selectively reference a speaker when they are uncertain about the referent of a label, rather than indiscriminantly looking at a speaker. In the present research, we capitalize on this method to ask whether preschool-aged children's social referencing is selective based on uncertainty generated by graded evidence.

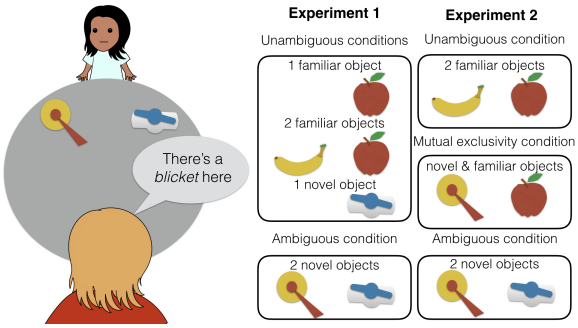

Figure 1: Study design for Experiments 1 and 2.

Experiment 1

The present research adapts the method of Vaish et al. to examine whether preschool-aged children's social referencing is sensitive to uncertainty based on referential ambiguity, and whether it reflects graded uncertainty about a word-object mapping. In Experiment 1, we examined whether children would reference a speaker more often when she produced a referentially ambiguous label. Children sat across from an experimenter who labeled an object on the table between them (Figure 1). Across trials, there were either one or two objects on the table, which were either familiar or unfamiliar to the child. We expected children to be uncertain about the label's referent only on trials with two unfamiliar objects, when the object-label mapping was not known and could not be inferred. After labeling an object, the experimenter asked the child to place the named object in a bucket.

We were interested in the amount of social referencing children exhibited across the trial. We considered four different phases of each trial based on the notion that children might expect different social information at different stages of the task. Specifically, we predicted that children might expect the speaker's gaze direction to be informative during the labeling itself, as speakers tend to look at objects they refer to. We predicted that later in the trial, as children reached for an object and placed it in the bucket, children might expect confirmation of the accuracy of their choice. We were interested in whether children would reference the speaker to a greater degree across these phases when referential ambiguity was high, which would indicate sensitivity to epistemic uncertainty.

Methods

Participants We recruited a planned sample of 80 children ages 2-5 years from the Children's Discovery Museum in San Jose, California[1]. The sample included 20 2-year-olds (mean age 31.71 months), 20 3-year-olds (mean age 42.65 months), 20 4-year-olds (mean age 55.85 months), and 20 5-year-olds (mean age 65.11 months). An additional 20 children participated but were removed from analyses because they heard English less than 75% of the time at home (n = 10), because they were unable to complete at least half of the trials in the task (n = 4), because of parental interference (n = 1), or due to experimenter or technical errors (n = 5).

Stimuli and Design Children were presented with one or two objects, heard a label, and were asked to put the labeled object in a bucket. Half of the objects were selected to be familiar to children (e.g., a cow) and half were selected to be novel (e.g., a nozzle). There were four trial types: one-familiar, one-novel, two-familiar, and two-novel. There were three trials of each type, for a total of twelve trials. Trial types were presented sequentially in an order that was counterbalanced across participants. The assignment of individual objects to trial types was counterbalanced. On familiar trials, the familiar label for the target object was used (e.g., "cow"). On novel trials, a novel label was used (e.g., "dawnoo").

The critical manipulation was of referential ambiguity; on one-familiar and two-familiar trials, there was no referential ambiguity, as children were expected to be certain about the objects and their labels. Similarly, on one-novel trials, children were expected to be certain about the label referent as there was only one option. However, on two-novel trials, the referent was ambiguous, as the novel label could apply to either novel object.

Procedure Throughout the study, the child sat at one end of a large circular table, and the experimenter stood at the opposite end. Each trial of the task proceeded as follows: the experimenter placed one or two objects on the left and/or right sides of the table, out of reach of the child so that the child could not interact with the toys during the labeling event. For one-object trials, the location of the object (left or right) alternated between trials. After placing the objects, the experimenter said "Hey look, there's a (target) here." The experi-

[1]Planned sample size, exclusion criteria, and analysis plan preregistered at `https://osf.io/y7mvt`

menter gazed at the center of the table rather than the object she was labeling because we wanted to preserve the referential ambiguity throughout the trial. The experimenter waited approximately two seconds (based on a visual metronome placed within view) before saying, "Can you put the (target) in the bucket?" She then pushed the object(s) forward within reach of the child, and placed a plastic bucket in the center of the table, also within reach of the child. Prior to the twelve experimental trials, there were two training trials: a one-familiar trial and a two-familiar trial, to acquaint the child with the procedure. A camera placed to the side of the experimenter captured the participant's face, so that looking behavior could be coded from video.

Coding procedure Videos were coded using DataVyu software (`http://datavyu.org`). First, each trial was divided into four temporal phases: a *label* phase, which began at the utterance of the target label and ended when the experimenter began to slide the objects, a *slide* phase, which encompassed the sliding of the objects into the child's reach, a *planning phase*, which began at the end of the slide and ended when the child touched an object, and a *response* phase, which began when the child touched an object and ended when the child released the object into the bucket. After onsets and offsets of these phases had been coded, the coder recorded the number of looks the child made to the experimenter during each phase. We opted to code the number of looks rather than the duration of looks because we felt that looks from the stimuli to the experimenter and vice versa might allow children to integrate social and nonsocial information to solve the problem of reference. A second coder coded the number of looks for a quarter of the trials for each participant to establish reliability. For Experiment 1, ##% of trials were given the same number of looks by both coders. For Experiment 2, ##% of trials were given the same number of looks by both coders.

Results and Discussion

Results of Experiment 1 are presented in Figure 2. To test our prediction that referential ambiguity (i.e., having two novel objects) would produce more social referencing, we fit mixed-effects linear regression models separately for each phase with the following structure: `number of looks ~ number of objects * familiarity * age in months + (number of objects + familiarity | SID)`. A single model with phase as a factor did not converge.

We did not find any main or interaction effects of number of objects, familiarity, or age on number of looks during the label or slide phases. However, we found an interaction effect of number of objects and familiarity during the planning ($\beta = 0.21$, $p < .001$) and response phases ($\beta = 0.6$, $p < .001$), such that 2-novel trials were associated with more looking. There was no interaction with age in either phase[2]. In summary, children looked to the experimenter more often when planning and executing a response under uncertainty. These results suggest that children were aware that they did not have sufficient knowledge to answer independently, and they attempted to resolve their uncertainty using social referencing.

We did not find the expected effect of referential ambiguity in the label phase. It is possible that children failed to predict that they would need more information until later in the trial, when they were actually faced with making a decision. Another possibility is that children's looking was at ceiling during the labeling phase, perhaps because children look at someone who is speaking regardless of the need for referential disambiguation. A third possibility is that this is an artifact of our design, in which the experimenter gazed at the center of the table rather than the referent of her label. Children may have realized that the experimenter's gaze direction during labeling was not informative. Experiment 2 tests this possibility and examines whether children's social referencing is sensitive to graded uncertainty.

Experiment 2

Experiment 2 was designed to replicate Experiment 1 and to investigate whether children's social referencing is sensitive to uncertainty based on graded evidence about a label's referent. Since we did not observe any difference between the one-familiar and one-novel trials, we eliminated single-object trials, leaving the 2-familiar and 2-novel trials. In addition, we added 1-novel-1-familiar trials ("mutual exclusivity trials"). For these trials, we expected that children would be able to infer the referent by excluding the familiar object as a possibility(Markman & Wachtel, 1988). We predicted that children might be less certain about their choice on these trials compared to when the label and referent were familiar to them (familiar trials), but more confident than when there are no cues to reference (novel trials).

In addition, we manipulated between participants whether or not the experimenter's gaze during labeling was informative (she gazed at either the referent of her label or the center of the table), allowing us to determine whether children selectively reference gaze during labeling when gaze is expected to be informative. The manipulation of informativity of gaze during labeling also meant that participants in the referential gaze condition had an additional referential cue, which might decrease uncertainty in the remainder of the trial.

In Experiment 1, we did not observe an effect of age on looking. Thus, we restricted the current sample to 3- and 4-year-olds.

Methods

Participants We recruited a planned sample of 80 children ages 3-4 years from the Children's Discovery Museum in San Jose, California[3]. The sample included 40 3-year-olds (mean age 42.89 months) and 40 4-year-olds (mean age 53.47 months). An additional 20 children participated but were removed from analyses because they heard English less than

[2] `https://github.com/emilyfae/socref_uncert`

[3] Planned sample size, exclusion criteria, and analysis plan preregistered at https://osf.io/y7mvt/

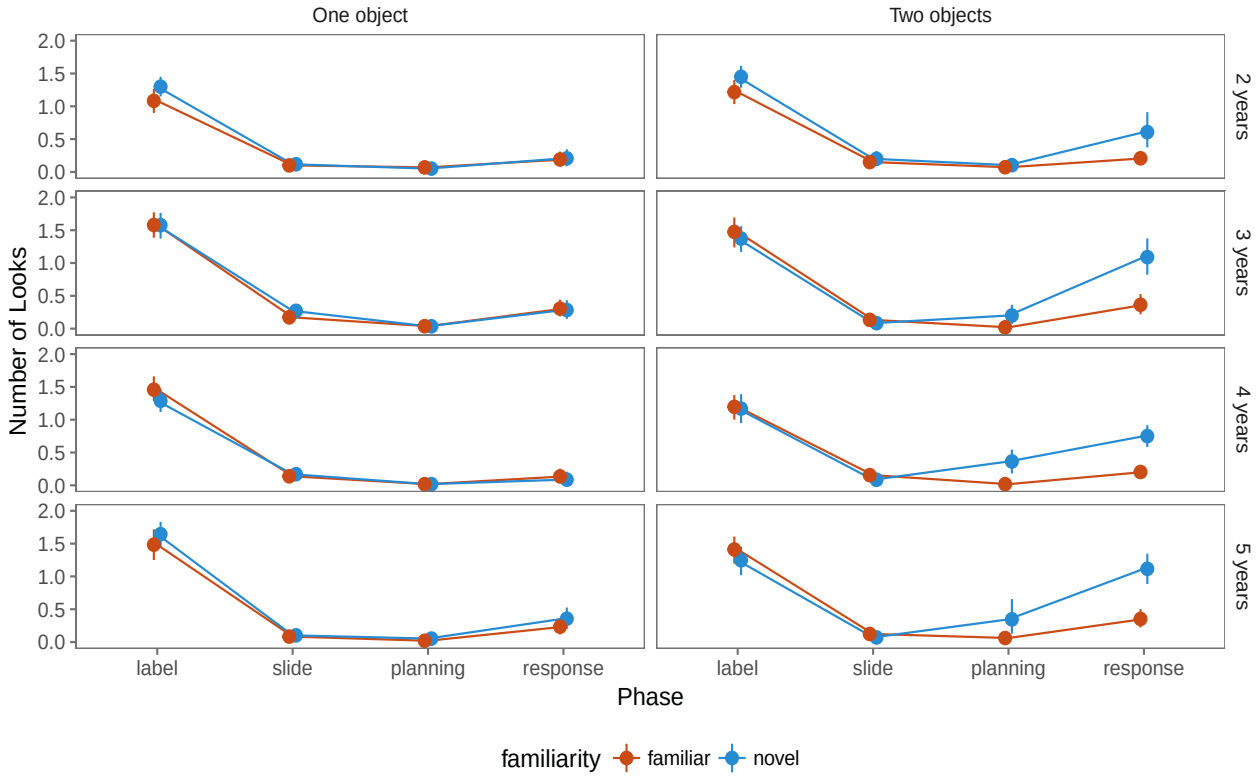

Figure 2: Results of Experiment 1. Number of looks to the experimenter across phases and conditions. Error bars are 95 percent confidence intervals.

75% of the time at home ($n = 9$), because they were unable to complete at least half of the trials in the task ($n = 7$), or due to experimenter or technical errors ($n = 4$).

Stimuli and Design The stimuli and design were similar to Experiment 1, except that we eliminated 1-object trials. Instead, we included three trial types: 2-familiar ("familiar"), 2-novel ("novel"), and 1-novel-1-familiar ("mutual exclusivity"). There were four of each trial type, totaling twelve trials. In addition, we manipulated the experimenter's gaze behavior between participants. For half of the participants, she looked at the center of the table on every trial; for the remaining half, she looked at the object she referred to on every trial.

Procedure The procedure was identical to Experiment 1, except that there were three practice trials rather than two, so that children could experience every trial type.

Results and Discussion

Results of Experiment 2 are presented in Figure 3. To quantify the main and interactive effects of familiarity, gaze informativity, phase, and age on social referencing, we fit a mixed-effects linear regression model with the following structure: `number of looks ~ familiarity * age in months * gaze * phase + (familiarity | SID)`. In contrast to Experiment 1, a model with phase as a predictor converged.

We were interested in several questions. First, do children reference a speaker more often when the objects and label are novel? Phase interacted with familiarity such that the response phase of novel trials was associated with significantly more looks ($\beta = 0.51$, $p < .001$). This is consistent with our finding from the linear model of the response phase in Experiment 1. However, in contrast to Experiment 1, we did not observe that looking was significantly greater for novel trials in the planning phase.

We were also interested in whether mutual exclusivity trials would elicit an intermediate amount of social referencing. We observed a three-way interaction of familiarity, gaze, and phase, such that the response phase of mutual exclusivity trials in the no-referential-gaze condition was associated with significantly more looks ($\beta = 0.39$, $p < .01$). In sum, mutual exclusivity trials were associated with greater looking during the response phase when the experimenter provided informative gaze compared to when she did not. This finding is intriguing given that children should be able to solve mutual exclusivity trials without gaze information. Instead, they remain relatively uncertain while executing a decision if excluding the familiar object is their only cue to reference, but this uncertainty is resolved if the experimenter gazes at the correct object during labeling. On the other hand, informative gaze during labeling did not lessen the amount of social referencing during the response phase for novel trials, suggesting that gaze information alone was not sufficient to reduce uncertainty. Children may require multiple cues

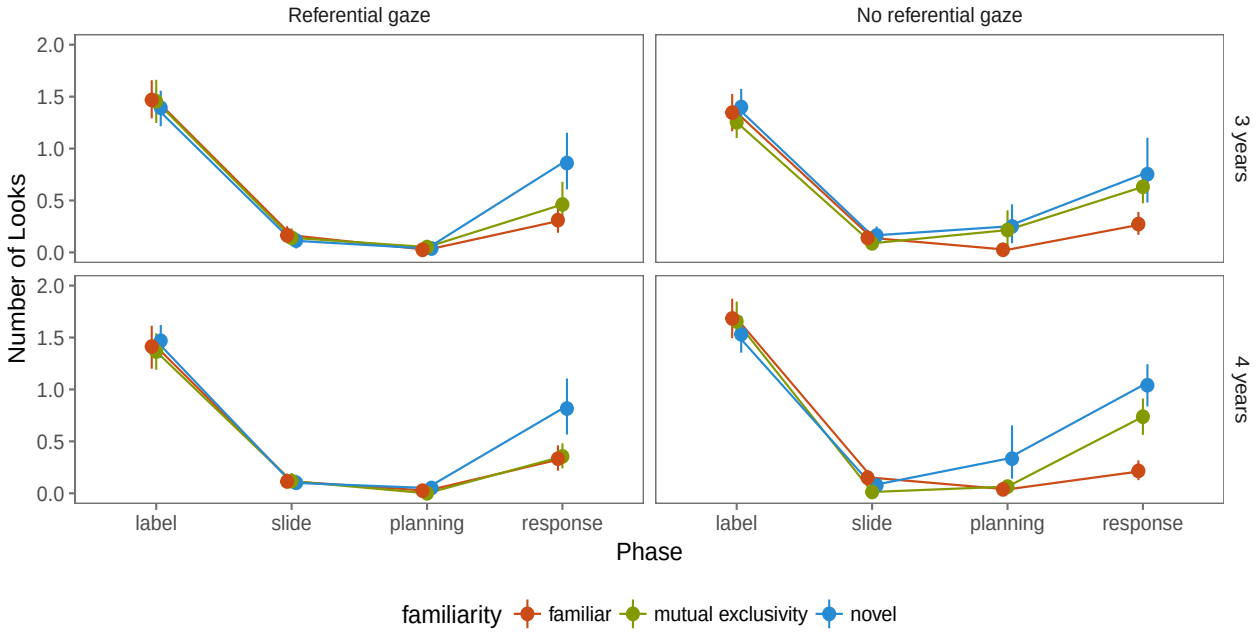

Figure 3: Results of Experiment 2. Number of looks to the experimenter across phases and conditions. Error bars are 95 percent confidence intervals.

This provides evidence for sensitivity to graded evidence.

Finally, we observed a four-way interaction such that the response phase of novel trials in the gaze condition was associated with more looking with increasing age ($\beta = 0.06$, $p < .01$), suggesting that children may become more selective in their social referencing as they get older. It may be that children improve in their ability to monitor the need for disambiguating information, or they may become more likely to recognize that social information can be a source of disambiguation.

Lastly, we again did not find selective social referencing during the label phase, even when referential gaze was available. This rules out the possibility that children were less selective during this phase because they learned that gaze direction was not informative.

General Discussion

Being able to detect uncertainty in mental representations seems critical for self-guided learning. During the preschool years, children are increasingly able to actively gather their own information through help-seeking and exploration. Do children monitor their own uncertainty to guide these behaviors, or are external features of the environment sufficient to guide learning (e.g., children might ask for help with a complex looking toy in response to perceptual features rather than epistemic states). Here, we examined whether children's social referencing in a word-learning scenario was calibrated to uncertainty about a label's referent.

We found strong evidence of selective social referencing under referential ambiguity when children were making their decision about which object was the target. In Experiment 1, we additionally found evidence for selective social referencing as children planned their decision, though this was not replicated in Experiment 2. We speculate that children may have referenced the speaker during the decision process because they expected confirmation of the accuracy of their choice, either implicitly through the adult's facial expressions, or through explicit feedback.

Importantly, we also found evidence for selectivity in social referencing based on graded uncertainty. In Experiment 2, we manipulated the amount of evidence available to children in two ways. First, we included mutual exclusivity trials consisting of a novel label paired with one familiar and one novel object, which we predicted might lead to intermediate levels of uncertainty compared to two-familiar and two-novel trials. Second, we manipulated whether or not the speaker's gaze direction was informative. We found that children treated mutual exclusivity trials more like familiar trials when they had received helpful gaze, but more like novel trials when they had not, suggesting that the combination of mutual exclusivity cues and gaze direction cues were required for children to feel confident about the object-label mapping. It is possible that they remain uncertain about a new label even after they have acquired it, if they have only heard it once and not received confirmation of its accuracy, for example, through gaze monitoring. Importantly, informative gaze during labeling did not lessen the amount of social referencing during the response phase for novel trials, suggesting that gaze information alone was not sufficient to reduce uncertainty.

On the other hand, we found no evidence for selectivity as the object was being labeled, or as the objects were being slid

into reach. One interpretation of this pattern of results is that preschool-aged children do not recognize the need for disambiguating information when a referent is ambiguous until they are in the position of needing to make a decision. However, another possibility is that children spontaneously look at a speaker regardless of the need for disambiguating information, and additional looking on top of this baseline was not needed or possible. A related possibility is that the labeling phase of our task was too short for children to produce extra looking on top of baseline in response to uncertainty. A follow-up to the present work that includes a longer labeling period or a greater reward tradeoff between attentional options would help to distinguish among these possibilities. Overall, these results provide evidence that preschool-aged children monitor graded uncertainty in their mental representations and act on that uncertainty through information-gathering behaviors.

Acknowledgements

We thank Veronica Cristiano for assisting with data collection.

References

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*(5), 875–890.

Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition*, *3*(4), 207–220.

Carpenter, M., Nagel, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development*, *63*(4), i–iii–v–vi–1–174.

Chouinard, M. M., Harris, P. L., & Maratsos, M. P. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, *72*, 1–129.

Coughlin, C., Hembacher, E., Lyons, K. E., & Ghetti, S. (2015). Introspection on uncertainty and judicious help-seeking during the preschool years. *Developmental Science*, *18*(6), 957–971.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata & M. H. Johnson (Eds.), *Processes of change in brain and cognitive development* (pp. 249–274). Oxford: Oxford University Press: Oxford University Press.

Destan, N., Hembacher, E., Ghetti, S., & Roebers, C. M. (2014). Early metacognitive abilities: The interplay of monitoring and control processes in 5- to 7-year-old children. *Journal of Experimental Child Psychology*, *126*(C), 213–228.

Flavell, J. H., Green, F. L., Flavell, E. R., Harris, P. L., & Astington, J. W. (1995). Young Children's Knowledge about Thinking. *Monographs of the Society for Research in Child Development*, i–iii–v–vi–1–113.

Ghetti, S., Lyons, K. E., Lazzarin, F., & Cornoldi, C. (2008). The development of metamemory monitoring during retrieval: The case of memory strength and memory absence. *Journal of Experimental Child Psychology*, *99*(3), 157–181.

Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they dont know. *Proceedings of the National Academy of Sciences*, *113*(13), 3492–3496.

Hembacher, E., & Ghetti, S. (2014). Dont Look at My Answer Subjective Uncertainty Underlies Preschoolers Exclusion of Their Least Accurate Memories. *Psychological Science*, *25*(9), 1–9.

Lipowski, S. L., Merriman, W. E., & Dunlosky, J. (2013). Preschoolers can make highly accurate judgments of learning. *Developmental Psychology*, *49*(8), 1505–1516.

Lyons, K. E., & Ghetti, S. (2013). I Don't Want to Pick! Introspection on Uncertainty Supports Early Strategic Behavior. *Child Development*, *84*(2), 726–736.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*, 121–157.

Metcalfe, J., & Finn, B. (2013). Metacognition and control of study choice in children. *Metacognition and Learning*, *8*(1), 19–46.

Paulus, M., Proust, J., & Sodian, B. (2013). Examining implicit metacognition in 3.5-year-old children: an eye-tracking and pupillometric study. *Frontiers in Psychology*, 1–7.

Robinson, M. D., Johnson, J. T., & Herndon, F. (1997). Reaction Time and Assessments of Cognitive Effort as Predictors of Eyewitness Memory Accuracy and Confidence. *Journal of Applied Psychology*, *82*, 416–425.

Schneider, W. (2008). The Development of Metacognitive Knowledge in Children and Adolescents: Major Trends and Implications for Education. *Mind, Brain, and Education*, *2*(3), 1–8.

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, *43*(4), 1045–1050.

Sodian, B., Thoermer, C., Kristen, S., & Perst, H. (2012). Metacognition in infants and young children. In M. J. Beran, J. Brandl, J. Perner, & J. Proust (Eds.), *Foundations of metacognition* (pp. 119–133).

Vaish, A., Demir, . E., & Baldwin, D. (2011). Thirteen- and 18-month-old Infants Recognize When They Need Referential Information. *Social Development*, *20*(3), 431–449.